EDA ATLANTA & LOS ANGELES

Atlanta Analysis: begins on page 1,

Los Angeles Analysis: begins on page 7

EDA Pumpkin prices Atlanta, Kaggle Data Set, Retail Store Costs

1. note: typo for origin, it is listed as orgin

2. Cost analysis of pumpkins in Atlanta area: in order to purchase pumpkins from different regions of the country.

3. No primary key or id in order to join tables from other parts of the country in this Kaggle dataset

--display data

SELECT \*

FROM atlanta;

Graphical user interface, text, application, email

Description automatically generated

--total count of each package

SELECT package, COUNT(package)

FROM atlanta

GROUP BY package

ORDER BY package;

Application, table

Description automatically generated

--pricing per package

--1: high price $18 low price $14.50

--2: high price $154.75 low price $123.50

--3: high price $286.00 low price $140.00

--high and low price per package code

SELECT high\_price, low\_price, package, rank

FROM

(SELECT high\_price, low\_price, package,

RANK () OVER(PARTITION BY package ORDER BY high\_price DESC) as rank

FROM atlanta)AS x

WHERE x.rank <2

SELECT variety, COUNT(variety) AS variety\_count

FROM atlanta

GROUP BY variety

ORDER BY variety;

Table

Description automatically generated

--pricing per variety

--1: high price $154.75 low price $140.00

--2: high price $18.00 low price $16.50

--3: high price $286.00 low price $14.50

--high and low price per variety code

SELECT high\_price, low\_price, variety, rank

FROM

(SELECT high\_price, low\_price, variety,

RANK () OVER(PARTITION BY variety ORDER BY low\_price ) as rank

FROM atlanta)AS x

WHERE x.rank <2

SELECT orgin, COUNT(orgin) AS orgin\_count

FROM atlanta

GROUP BY orgin

ORDER BY orgin;

Graphical user interface, application

Description automatically generated

--pricing per orgin

--1: high price $145.00 low price $140.00

--2: high price $150.00 low price $150.00

--3: high price $286.00 low price $14.50

--4: high price $135.00 low price $123.50

--5: high price $140.00 low price $16.50

--high and low price per orgin code

SELECT high\_price, low\_price, orgin, rank

FROM

(SELECT high\_price, low\_price, orgin,

RANK () OVER(PARTITION BY orgin ORDER BY high\_price DESC) as rank

FROM atlanta)AS x

WHERE x.rank <2

SELECT item\_size, COUNT(item\_size) AS item\_size\_count

FROM atlanta

GROUP BY item\_size

ORDER BY item\_size;

Graphical user interface, application, table

Description automatically generated

--pricing per item\_size

--1: high price $154.75 low price $140.00

--2: high price $145.00 low price $140.00

--3: high price $150.00 low price $140.00

--4: high price $154.75 low price $150.00

--5: high price $286.00 low price $14.50

--6: high price $154.75 low price $145.00

--high and low price per item\_size code, see above

SELECT MIN(low\_price) AS min\_low, MAX(low\_price) AS max\_low, ROUND(AVG(low\_price), 2) AS avg\_low,

MIN(high\_price) AS min\_high, MAX(high\_price) AS max\_high, ROUND(AVG(high\_price),2) AS avg\_high

FROM atlanta;

Graphical user interface, text, application

Description automatically generated

SELECT low\_price, high\_price, (high\_price-low\_price) AS price\_diff

FROM atlanta

GROUP BY price\_diff, high\_price, low\_price

ORDER BY price\_diff;

Table

Description automatically generated

SELECT date, COUNT(date) AS date\_count

FROM atlanta

GROUP BY date

ORDER BY date;

Graphical user interface, application, table

Description automatically generated

--low\_price by dates ordered by lowest price, best date to purchase pumpkins from Atlanta

SELECT low\_price, date

FROM atlanta

GROUP BY low\_price, date

ORDER BY low\_price;

Table

Description automatically generated

--price of pumpkins: high and low prices by date:

--partition by low\_price, ordered by date

Table

Description automatically generated with low confidence

--partition by date, ordered by low\_price

Table

Description automatically generated with medium confidence

--partition by high\_price, ordered by date

Table

Description automatically generated with low confidence

--partition by date, orered by high\_price

Table

Description automatically generated with medium confidence

EDA Pumpkin prices los\_angeles, Kaggle Data Set, Retail Store Costs

1. Cost analysis of pumpkins in Los Angeles area: in order to purchase pumpkins from different regions of the country.

2. No primary key or id in order to join tables from other parts of the country in this Kaggle dataset

--display data

Application

Description automatically generated with medium confidence

Text

Description automatically generated

--pricing per package

--1: high price $130.00 low price $0.24

--2: high price $30.00 low price $24.50

--3: high price $375.00 low price $0.30

--4: high price $150.00 low price $125.00

--5: high price $250.00 low price $0.55

--There is a large price disparity for some of the packages, reason unknown

--high and low price per package code

Graphical user interface, text, application

Description automatically generated

Graphical user interface, application

Description automatically generated

--pricing per variety

--1: high price $250.00 low price $0.50

--2: high price $250.00 low price $0.30

--3: high price $250.00 low price $0.24

--4: high price $375.00 low price $80.00

--5: high price $250.00 low price $0.50

--6: high price $ 24.50 low price $24.50

--7: high price $30.00 low price $30.00

--high and low price per variety code

--high and low prices per variety

SELECT high\_price, low\_price, variety, rank

FROM

(SELECT high\_price, low\_price, variety,

RANK () OVER(PARTITION BY variety ORDER BY low\_price ) as rank

FROM los\_angeles)AS x

WHERE x.rank <2

--all pumpkins originate only from California

Graphical user interface, text, application

Description automatically generated

--pricing per origin

--1: high price $375.00 low price $0.24

--high and low price per origin code

A picture containing text

Description automatically generated

Graphical user interface, table

Description automatically generated

--pricing per item\_size

--1: high price $250.00 low price $80.00

--2: high price $250.00 low price $0.50

--3: high price $180.00 low price $0.30

--4: high price $375.00 low price $0.55

--high and low price per item\_size code, see above

--minimum, maximum, and average prices for low and high price points

Graphical user interface, text, application

Description automatically generated

--low, high, and the price difference grouped and then ordered in descending order

Table

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

--highest priced days to purchase pumpkins

Table

Description automatically generated

--price of pumpkins: high and low prices by date:

--partition by low\_price, ordered by date

--best dates to order pumpkins

A picture containing table

Description automatically generated

--partition by high\_price, ordered by date

--most expensive dates to order pumpkins

Table

Description automatically generated